

An evaluation of Inherent Optical Property data for inclusion in the NASA bio-Optical Marine Algorithm Data set

P. Jeremy Werdell
NASA Ocean Biology Processing Group
Science Systems and Applications, Inc.

Document Version 1.1, corresponding to NOMAD Version 1.3
19 September 2005

The original version of the NASA bio-Optical Marine Algorithm Data set (NOMAD; Werdell and Bailey 2005; <http://seabass.gsfc.nasa.gov/cgi-bin/nomad.cgi>) includes 3,467 coincident observations of spectral water-leaving radiances, $L_w(\lambda)$, surface irradiances, $E_s(\lambda)$, and diffuse downwelling attenuation coefficients, $K_d(\lambda)$, encompassing chlorophyll *a* concentrations, C_a , ranging from 0.012 to 72.12 m mg^{-3} . In terms of ocean color algorithm development, its lack of coincident observations of inherent optical properties (IOPs), such as spectral absorption and backscattering coefficients, limits the utility of NOMAD to the development of empirical, reflectance-based C_a algorithms (e.g., O'Reilly et al. 2000). The next generation of remote-sensing algorithms, such as multi-component semi-analytical algorithms (e.g., Roesler and Perry 1995, Hoge and Lyon 1996, Garver and Siegel 1997, Carder et al. 1999, Lee et al. 2002), requires IOPs, most often spectral absorption coefficients for phytoplankton, $a_{ph}(\lambda)$, non-algal detrital material, $a_d(\lambda)$, and dissolved organic material, $a_g(\lambda)$, and spectral backscattering coefficients for particles, $b_{pp}(\lambda)$. As such, in July 2005, the NASA Ocean Biology Processing Group (OBPG) commenced adding IOPs to NOMAD. Various components of this effort are described in this report. As of this writing, NOMAD includes absorption coefficients derived from laboratory spectroscopy and backscattering coefficients estimated from *in situ* profiling instrumentation.

I. Data acquisition

All data were acquired from the NASA SeaWiFS Bio-optical Archive and Storage System (SeaBASS; Hooker et al. 1994, Werdell and Bailey 2002). Data were limited to those collected following OBPG-defined protocols (Pegau et al. 2003) to ensure consistency amongst data contributed by multiple investigators and to, therefore, facilitate the evaluation of a large volume of data. Please refer to Section 2.1 of Werdell and Bailey (2005) for a detailed description of data acquisition methods, contributing Principal Investigators, and global distribution. For the remainder of this report, spectral dependence is implied and no longer reported.

II. Absorption coefficients

We acquired approximately 6,400 observations of spectral absorption coefficients derived exclusively by spectroscopy. Typically, data products included a_d , a_g , and the absorption coefficient for total particulate matter, a_p . Most semi-analytical algorithms require a_{ph} , which is calculated via $a_p - a_d$. For storage efficiency, and as a_{ph} is a derived product, we retained only a_p , a_d , and a_g . On occasion, a_{ph} was recorded *in lieu* of a_p . Each time this occurred, a_p was reconstructed via $a_{ph} + a_d$. To expedite organizational and processing efficiency, we consolidated all coincident spectra into single data files (one file per depth per station), if not available from SeaBASS in such a format. Observations collected below 30-meters were discarded, as they were not optically significant for this exercise. All remaining spectra were visually

inspected and geophysically unreasonable data (e.g., those with excessive noise or monotonically increasing magnitudes for a_d and a_g) were removed.

To remove moderate noise, often resulting from instrument artifacts or poor sample baselines, we derived smooth fits for a_d and a_g following the form of Roesler et al. (1989):

$$a_x(\lambda) = a_x(\lambda_0) \exp [-S_x (\lambda - \lambda_0)] \quad (1)$$

where x indicates either d or g , S defines the spectral shape of the curves, and λ_0 is a reference wavelength, often 400-nm. For each sample, we calculated average values for S via linear least-squares regression over the ranges 380 – 530-nm and 380 – 600-nm. The fit with the higher correlation coefficient was retained. All original spectra and fits were simultaneously visually inspected and data with fitting errors were reanalyzed or discarded.

We reduced both original and fit spectra to the twenty nominal wavebands listed in Table 6 of Werdell and Bailey (2005). For both a_d and a_g , data at specific wavebands were discarded unless the following condition was satisfied: $0.5 \leq [a_x(\lambda) / \hat{a}_x(\lambda)] \leq 2$, where \hat{a}_x indicates fit data. For a_d , these outliers were typically located near the near-infrared C_a absorption peak, likely resulting from an incomplete methanol extraction. For a_g , such outliers showed no spectral dependence and were often suspected to result from cell lyses during filtering.

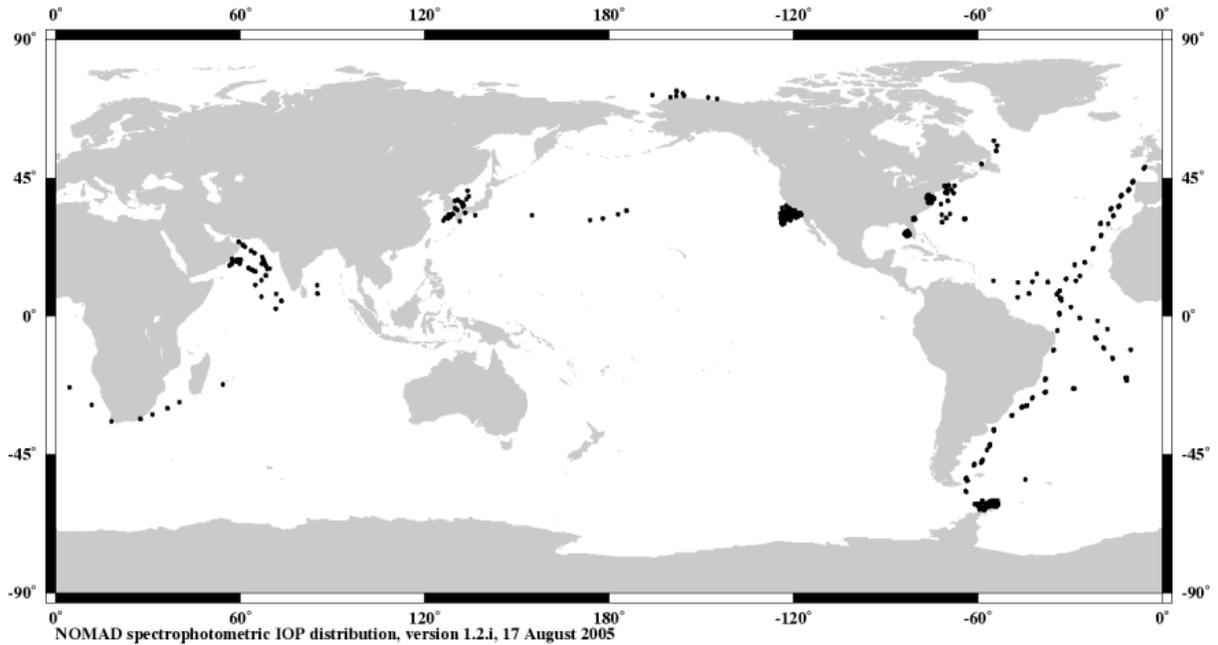


Figure 1. Global map of spectrophotometric absorption data included in NOMAD.

For stations with observations at multiple depths, data were optically weighted following Section 2.3.2 of Werdell and Bailey (2005). Briefly, we used the method of Gordon and Clark (1980) and relevant $K_d(490)$ measurements to derive a single remote-sensing relevant absorption spectrum for each station. To limit the volume of noisy data included in the weighting scheme, only a_p , \hat{a}_d , and \hat{a}_g were considered. When profiles were unavailable, the surface-most spectrum was selected if collected at a depth less than 5-meters. A NOMAD binary flag (bit 20; $2^{20} = 1,048,576$; see Table 4 of Werdell and Bailey (2005) for details) was set for each optically weighted station. Approximately two-thirds of the input stations

included observations from multiple depths, with 3.5 discrete samples collected on average over the top 30-meters of the water column.

We visually and statistically evaluated both a_{ph} and \hat{a}_{ph} ($= a_p - \hat{a}_d$). Spectra with negative values or questionable shapes were reanalyzed or discarded. As a final quality control measure, we inspected time series of derived S_x and $a_x(443)$ on a cruise-by-cruise basis to identify local outliers. In the end, 883 NOMAD records were updated with a_p , \hat{a}_d , and \hat{a}_g (Figure 1, Table 1). We used temporal and spatial thresholds of ± 3 hours and 0.1° , respectively, to merge these data with the original 3,467 NOMAD stations (see Section 2.5 of Werdell and Bailey (2005) for a detailed discussion of this process). Increasing the temporal threshold to 4 or 5 hours would have added only 3 or 5 stations, respectively. Absorption data are available online as of NOMAD Version 1.2. To facilitate community evaluation of our processing, the surface-most original (unmodified) spectra, a_p , a_d , and a_g , and surface-most fit spectra, \hat{a}_d and \hat{a}_g , for all 883 stations are available online separately in a secondary data file.

data products	number
AOPs + $C_a + K_d$ (= baseline)	3,467
baseline + spectrophotometric a	883
baseline + b_b	184
baseline + spectrophotometric $a + b_b$	97

Table 1. The quantity of coincident data products in NOMAD.

III. Backscattering coefficients

We acquired nearly 1,200 measurements of the spectral backscattering coefficient, b_b , obtained using HOBI Labs HydroScat, a -beta, and c -beta sensors (<http://www.hobilabs.com>), WET Labs ECO-BB and ECO-VSF sensors (<http://www.wetlabs.com>), and Wyatt Technology Corporation DAWN photometers (<http://www.wyatt.com>). All data were collected as vertical depth profiles with the exception of those from the Scotia Prince Ferry program, which were collected underway via a fixed-depth shipboard flow-through system (Balch et al. 2003). In processing the data, all contributors applied a sigma-correction to correct for light attenuation in the path of the instrument, and most used the dimensionless coefficient, χ_p , derived by Maffione and Dana (1997) to relate the volume scattering function to b_b . For the latter, data from the Plumes and Blooms program were processed using the χ_p from Boss and Pegau (2001), and Scotia Prince Ferry data collected in 2003 and later were processed using χ_p from Vaillancourt et al. (2004). From a semi-analytical algorithm perspective, the target product is b_{bp} , however, we opted not to subtract the backscattering contribution of seawater, b_{bw} , from b_b ($b_{bp} = b_b - b_{bw}$), as b_{bw} is well known and readily available to an end user (Morel 1974; Lee et al. 1996).

We visually inspected all depth profiles and removed those with significant noise or without measurements collected more shallow than 5-meters. Nearly 80% of the profiled data were binned to 1-meter depth resolution prior to submission to SeaBASS. We subsequently binned the remaining profiles via: (i) designation of a site-specific bin size (e.g., 1-meter); (ii) application of a statistical filter to data within each bin to exclude observations outside $M_z \pm 1.5 s_z$, where M_z and s_z are the population median and standard deviation of the bin; and (iii) calculation of the arithmetic mean of all remaining data points. As in Section II, profiled data were optically weighted following Section 2.3.2 of Werdell and Bailey (2005). Briefly, we used the method of Gordon and Clark (1980) and relevant $K_d(490)$ measurements to derive a single remote-sensing relevant backscattering spectrum for each station. A NOMAD binary flag (bit 21; $2^{21} = 2,097,152$; see Table 4 of Werdell and Bailey (2005) for details) was set for each optically weighted station. We processed the remaining Scotia Prince Ferry flow-through data following Section 2.3.3 of Werdell and Bailey (2005). These data required an additional magnitude adjustment as described in Balch

et al. (2003). All spectra were visually inspected and geophysically unreasonable data (e.g., those with monotonically increasing magnitudes or negative values) were removed.

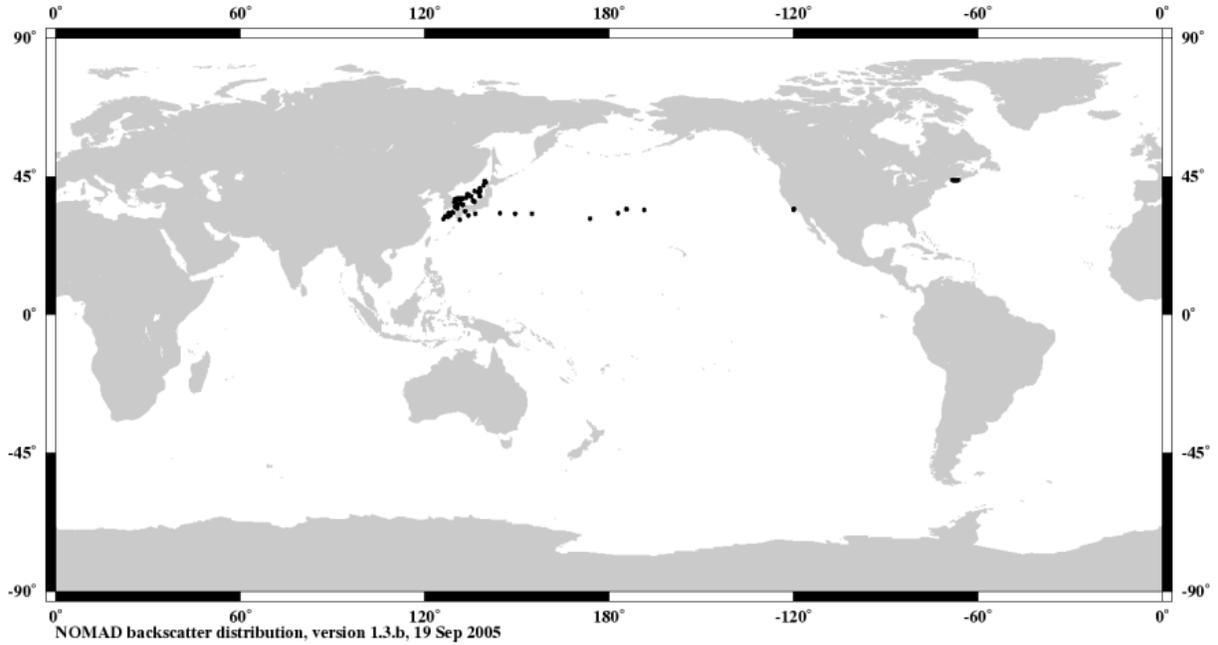


Figure 2. Global map of backscattering data included in NOMAD.

To remove moderate noise, often resulting from instrument artifacts or calibration, we derived smooth fits for b_b presupposing the form:

$$b_b(\lambda) = b_{bw}(\lambda) + \{b_{bp}(\lambda_0) \{\lambda/\lambda_0\}^\nu\} \quad (2)$$

where ν is a unitless parameter that defines spectral slope of particulate backscattering (Morel 1973), for example, -1.7 and -0.3 for small and large particles, respectively (Kopelevich 1983), and λ_0 is a reference wavelength, often 550-nm. For reference, the spectral slope is approximately -4.3 for molecular backscattering (Morel 1974). For each sample, we calculated average values for the spectral slope, $\hat{\nu}$, over the range 380 – 700-nm via nonlinear multidimensional minimization of the bracketed part of (2) onto the measured b_{bp} (Press et al. 1992). Values ranged from -2.46 to 0.0 , with a mean of -1.02 . Only 11 observations were ≤ -2.0 . We reconstructed b_b using (2) at the twenty nominal wavebands listed in Table 6 of Werdell and Bailey (2005) using $\hat{\nu}$ and the calculated regression intercept. All original spectra and fits (for both b_b and b_{bp}) were simultaneously visually inspected and data with fitting errors were reanalyzed or discarded.

In the end, 184 NOMAD records were updated with b_b (Figure 2, Table 1). We used temporal and spatial thresholds of ± 3 hours and 0.1° , respectively, to merge these data with the original 3,467 NOMAD stations (see Section 2.5 of Werdell and Bailey (2005) for a detailed discussion of this process). Increasing the temporal threshold to 4 or 5 hours would not have included additional stations. Backscattering data are available online as of NOMAD Version 1.3. To facilitate community evaluation of our processing, the original (unmodified) spectra for all 200 stations, including both the adjusted and unadjusted Scotia Prince Ferry data, are available online separately in a secondary data file.

We recommend that end users adopt cautious approaches to using these data. First, measuring backscattering is complicated and multiple approaches and instruments have been used to estimate b_b . Contributors to SeaBASS already reprocess their b_b profiles regularly, and with time, further improved techniques will be realized. Second, as additional data become available, for example, total scattering and attenuation coefficients, subsequent data products will be calculable, such as the backscattering efficiency factor, and OBPG processing techniques will evolve (see, e.g., Roesler and Boss 2003). Finally, we hope to receive community feedback on these methods, which may result in a NOMAD reprocessing based on community recommendations. Ultimately, we expect updates to NOMAD b_b to be somewhat common.

IV. Acknowledgements

Thanks to S. Bailey, T. Moore, S. Maritorena, and D. Court for advice on NOMAD data evaluation. Thanks to W. Balch, B. Bowler, N. Guillocheau, M. Kahru, T. Kostadinov, and N. Nelson for assistance in processing and reviewing their data.

V. References

- Balch, W.M. and Drapeau, D.T. (2003), Backscattering by coccolithophorids and coccoliths: sample preparation, measurement, and analysis protocols. In J.L. Mueller, G.S. Fargion, and C.R. McClain (Eds.), *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 5, Volume V: Biogeochemical and Bio-optical Measurements and Data Analysis Protocols* (pp. 27-36). NASA Tech. Memo. 2003-211621, Rev. 5, Vol. V, Greenbelt: NASA Goddard Space Flight Center.
- Boss, E., and Pegau, W.S. (2001), Relationship of light scattering at an angle in the backward direction to the backscattering coefficient. *Applied Optics*, 40, 5503-5507.
- Carder, K.L., Lee, Z.P., Hawes, S.K., and Kamykowski, D. (1999), Semianalytic Moderate-Resolution Imaging Spectrometer algorithms for chlorophyll *a* and absorption with bio-optical domains based on nitrate-depletion temperatures. *Journal of Geophysical Research*, 104, 5403-5421.
- Garver, S.A. and Siegel, D.A. (1997), Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation .1. Time series from the Sargasso Sea. *Journal of Geophysical Research*, 102, 18607-18625.
- Gordon, H.R. and Clark, D.K. (1980), Remote sensing optical properties of a stratified ocean: and improved interpretation. *Applied Optics*, 19, 3428-3430.
- Hoge, F.E. and Lyon, P.E. (1996), Satellite retrieval of inherent optical properties by linear matrix inversion of oceanic radiance models: An analysis of model and radiance measurement errors. *Journal of Geophysical Research*, 101, 16631-16648.
- Hooker, S.B., McClain, C.R., Firestone, J.K., Westphal, T.L., Yeh, E.N., and Geo Y. (1994), *The SeaWiFS Bio-optical Archive and Storage System (SeaBASS), Part 1*. NASA Tech. Memo. 104566, Vol. 20. Greenbelt: NASA Goddard Space Flight Center, pp. 37.
- Kopelevich, O.V. (1983), Small-parameter model of optical properties of sea water. In: A.S. Monin (Ed.), *Ocean Optics, I: Physical Ocean Optics* (pp. 208-234). Nauka Publishing, Moscow, (in Russian).

- Lee, Z.P., Carder, K.L., Peacock, T.G., Davis, C.O., and Mueller, J.L. (1996), Method to derive ocean absorption coefficients from remote-sensing reflectance. *Applied Optics*, 35, 453-462.
- Lee, Z.P., Carder, K.L., and Arnone, R.A. (2002), Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, 41, 5755-5772.
- Maffione, R.A. and Dana, D.R. (1997), Instruments and methods for measuring the backward-scattering coefficient of ocean waters. *Applied Optics*, 36, 6057-6067.
- Morel, A. (1973), Diffusion de la lumière par les eaux de mer: Résultats expérimentaux et approche théorique. In *NATO AGARD Lecture Series 61* (pp. 3.1.1-3.1.76).
- Morel, A. (1974), Optical properties of pure water and pure sea water. In: N.G. Jerlov and E. Steemann Nielsen (Eds.), *Optical Aspects of Oceanography* (pp. 1-24). Academic Press, London.
- O'Reilly, J.E., and 24 co-authors (2000), Ocean color chlorophyll *a* algorithms for SeaWiFS, OC2, and OC4: Version 4. In S.B. Hooker and E.R. Firestone (Eds.), *SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3* (pp. 9-23). NASA Tech. Memo. 2000-206892, Vol. 11, Greenbelt: NASA Goddard Space Flight Center.
- Pegau, S., Zaneveld, J.R.V., Mitchell, B.G., Mueller, J.L., Kahru, M., Wieland, J., and Stramska, M. (2003): *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume IV: Inherent Optical Properties: Instruments, Characterizations, Field Measurements and Data Analysis Protocols*. NASA Tech. Memo. 2003-211621, Rev. 4, Vol. IV, Greenbelt: NASA Goddard Space Flight Center, pp. 76.
- Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterlin, W.T. (1992), *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, London, pp 1020.
- Roesler, C.S. and Boss, E. (2003), Spectral beam attenuation coefficient retrieved from ocean color inversion. *Geophysical Research Letters*, 30, doi:10.1029/2002GL016185.
- Roesler, C.S., Perry, M.J., and Carder, K.L. (1989), Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters. *Limnology and Oceanography*, 34, 1510-1523.
- Roesler, C.S. and Perry, M.J. (1995), In-situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance. *Journal of Geophysical Research*, 100, 13279-13294.
- Vaillancourt, R.D., Brown, C.W., Guillard, R.R., and Balch, W.M. (2004), Light backscattering properties of marine phytoplankton: relationship to cell size, chemical composition and taxonomy. *Journal of Plankton Research*, 26, 191-212.
- Werdell, P.J. and Bailey, S.W. (2002), *The SeaWiFS Bio-optical Archive and Storage System (SeaBASS): Current Architecture and Implementation*. NASA Tech. Memo. 2002-211617, Greenbelt: NASA Goddard Space Flight Center, pp. 45.
- Werdell, P.J. and S.W. Bailey (2005), An improved bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, 98, 122-140.